University of Wisconsin - Madison

MAD/PH/849
UCD-94-34
RAL-94-099

September 1994

Further ways to distinguish single-lepton top quark signals from background at the Tevatron

V. Barger and E. Mirkes

Physics Department, University of Wisconsin, Madison, WI 53706, USA

J. Ohnemus

Physics Department, University of California, Davis, CA 95616, USA

R.J.N. Phillips

Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK

Abstract

Further kinematical variables are suggested, in which to compare the putative leptonic W plus 4-jet top quark signal with the QCD background. We show that the lepton rapidity asymmetry, the p_T-ranking of the tagged b-jet, the double-tag probability, the reconstructed $t+\bar{t}$ invariant mass, and the lepton energy in the parent top quark rest-frame, all display interesting differences between signal and background.

Evidence for top quark production in $p\bar{p}$ collisions at the Fermilab Tevatron collider was recently presented by the CDF collaboration [1], corresponding to $t\bar{t}$ production with $t \to bW^+$ and $\bar{t} \to \bar{b}W^-$; the D0 collaboration has also presented a top quark candidate event [2]. In the present work we address the b-tagged single-lepton signal [3], where one of the W bosons decays leptonically $W \to \ell\nu$ and the other decays hadronically to two jets $W \to jj$, while at least one of the four final jets is tagged as a probable b-jet by vertex or secondary-lepton criteria. The most serious background to this signal comes from QCD production of $W + 4$ jets; the CDF event criteria were designed to suppress this background, but nevertheless it will be desirable to make further cross-checks to establish as firmly as possible the credentials of the putative top quark signal in present and future data. The CDF collaboration has already presented several such supplementary checks, notably against the various jet transverse energy E_T distributions and correlations and against the summed-$E_T(j)$ distribution (first advocated in Ref. [4]).

In the present Letter we draw attention to five further quantities, where the behaviour of top quark signal and background are significantly different, and illustrate these differences by specific calculations. The quantities proposed here are:

(a) The forward-backward asymmetry in the charged-lepton rapidity distribution,

$$A(y_\ell) = \pm[d\sigma/dy(y_\ell) - d\sigma/dy(-y_\ell)]/[d\sigma/dy(y_\ell) + d\sigma/dy(-y_\ell)]\,, \qquad (1)$$

where the $\pm$ sign is equal to the lepton charge and $y > 0$ is the hemisphere in the p beam direction.

(b) The ranking-order (1,2,3,4) of the b-tagged jets (where the jets are p_T-ordered and jet 1 has highest p_T).

(c) The probability that a tagged event has two tagged jets.

(d) The invariant mass $m(t\bar{t})$ of the two top quarks, extracted by explicitly reconstructing the event from the lepton and jet momenta and the missing-E_T vector.

(e) The lepton energy E_ℓ in the reconstructed rest-frame of the semileptonically decaying top quark.

In the following, we shall discuss the qualitative features that cause signal and background to differ, and give quantitative examples. But first we describe our methods of calculation.

For our illustrations, we make parton-level Monte Carlo calculations. We generally calculate $t\bar{t}$ production at lowest order (α_s^2), with full spin correlations [5]

in the subsequent W boson decays, using the MRS set D0 parton distributions [6] at scale $Q = m_t$, taking $m_t = 174$ GeV [1] and normalizing to the next-to-leading-order total $t\bar{t}$ cross section [7]. However, for the lepton asymmetry, which vanishes at lowest order, we calculate $2 \to 3$ parton subprocesses at order α_s^3 [8], using the truncated shower approximation [9]. We calculate $W + 4$-jet production at leading order with the VECBOS program [10], supplemented by the programs developed in Ref. [11], at scale $Q = <p_T(j)>$ (the mean jet transverse momentum), with the same parton distributions and 4 quark flavors. We set out to impose approximately the same acceptance cuts as in the CDF single-lepton top quark search [1]. We regard the final four partons as jets, and require the transverse momenta $p_T(j)$, pseudorapidities $\eta(j)$, and separations $\Delta R(jj)$ of the first three jets to satisfy

$$p_T(j) > 20\,\mathrm{GeV}, \quad |\eta(j)| < 2.0, \quad \Delta R(jj) > 0.7, \quad (j = 1, 2, 3), \tag{2}$$

where $\eta = \ln(\tan\theta/2)$ and $(\Delta R)^2 = (\Delta\eta)^2 + (\Delta\phi)^2$; θ and ϕ are the usual polar and azimuthal angles measured with respect to the antiproton beam direction. These conditions are relaxed for the fourth jet:

$$p_T(4) > 13\,\mathrm{GeV}, \quad |\eta(4)| < 2.4, \quad \Delta R(j4) > 0.7. \tag{3}$$

In the neighborhood of these cuts, studies of jet fragmentation and instrumental effects have shown that the initiating parton p_T exceeds the measured jet E_T by about 5 GeV in the CDF apparatus [1]; thus the parton cuts above are intended to approximate the jet cuts $E_T(j = 1, 2, 3) > 15$ GeV and $E_T(4) > 8$ GeV in the CDF top quark analysis [1]. For leptons we require

$$p_T(\ell) > 20\,\mathrm{GeV}, \quad |\eta(l)| < 1.0, \quad \Delta R(\ell j) > 0.4, \quad (j = 1, 2, 3, 4), \tag{4}$$

where the lepton-jet separation approximates the lepton isolation criterion. We apply realistic gaussian smearing factors [12] to all parton and lepton momenta, to represent measurement errors, with CDF resolution values [13]. The missing-transverse energy vector $\not{E}_T$ is defined to be the negative sum of all the lepton and parton transverse energy vectors, and must satisfy

$$\not{E}_T > 20\,\mathrm{GeV}. \tag{5}$$

With these cuts, we calculate the total lepton-plus-four-jet signal from $t\bar{t}$ production to be 0.50 pb, to be compared with 0.062 pb from $Wbbjj$ alone and 2.3 pb from all $Wjjjj$ channels. (If we choose scale $Q^2 = <p_T(j)>^2 + M_W^2$ instead, these background numbers become 0.029 pb and 1.1 pb, respectively.) The signal/background ratio can then be improved by b-tagging.

In the CDF experiment, the efficiency for tagging one or more b-jets in a $t\bar{t}$ event is about 0.33 (combining vertex and lepton tagging approaches), corresponding to a probability $\epsilon_b \simeq 0.18$ per b-jet; the probability of a fake b-tag appears to be $\epsilon_q \simeq \epsilon_g \simeq 0.01$ per light-quark or gluon jet. We assume a probability

$\epsilon_c \simeq 0.05$ for a bogus c-jet tag. Tagging efficiencies are taken to be approximately independent of p_T and η. The cross section for each final configuration is multiplied by the corresponding probability that at least one of the jets is tagged; *e.g.*, the tagged cross sections for $Wgggg$, $Wc\bar{c}gg$, $Wb\bar{b}qq'$ production contain tag-probability factors 0.04, 0.12, 0.34, respectively. With these factors, Ref. [11] shows that the b-tagged $W+4$-jet background comes mainly from $W+4q/g$ final states, *i.e.*, from mistagging light-quark or gluon jets; next, in descending order of importance, come $W+2b+2q/g$, $W+c+3q/g$, and $W+2c+2q/g$ final states. The tagged $W+4$-jet background is therefore dominated by the $Wjjjj$ and $Wbbjj$ configurations, which can be calculated from the VECBOS program [10] alone. With these tag-factors, the net tagged $t\bar{t}$ single-lepton signal becomes 0.13 pb, to be compared with 0.055 pb (for scale $Q = <p_T(j)>$) from the combined $Wjjjj$ channels.

We now discuss and illustrate the kinematical distributions (a) – (e) described above, for the single-lepton $t\bar{t}$ signal and the $Wjjjj$ background at the Tevatron.

(a) Lepton asymmetry

For $Wjjjj$ production, two different effects contribute (in opposite directions) to a forward/backward lepton asymmetry. The preponderance of valence u-quarks over d-quarks in the proton causes W^+ (W^-) bosons to be produced preferentially in the p ($\bar{p}$) beam hemisphere. The $V-A$ couplings in the subprocess $ud \to W \to \ell\nu$ produce $\ell^+(\ell^-)$ preferentially along the $\bar{p}$ (p) axis in the W boson rest-frame. At the Tevatron energy the former effect prevails and the background $\ell^+(\ell^-)$ is predicted to be produced preferentially in the p ($\bar{p}$) beam hemisphere in the lab frame; see Fig. 1. On the other hand, the $p\bar{p} \to t\bar{t}$ tree-level mechanisms produce unpolarized top quarks, with no forward/backward asymmetry at lowest order; hence there is no lepton asymmetry here at order α_s^2. At order α_s^3, however, the subprocess $q\bar{q} \to t\bar{t}g$ (which dominates over other $2 \to 3$ subprocesses at the Tevatron energy for $m_t = 174$ GeV) produces t-quarks preferentially in the direction of the incoming antiquark $\bar{q}$, due to interference between gluons emitted from initial and final quarks [14]; this leads to a small lepton asymmetry of *opposite* sign to the $Wjjjj$ case. Note that the $Wjjjj$ effect arises from the isospin asymmetry within the valence quarks in p and $\bar{p}$, whereas the $t\bar{t}$ effect comes from the difference between the triplet and antitriplet color representations of valence quarks and antiquarks. The asymmetries and the underlying lepton rapidity distributions are shown in Figs. 1(a) and 1(b), respectively; both these quantities differ significantly between signal and background. If we define the forward lepton hemisphere as $y(\ell) > 0$ ($y(\ell) < 0$) for ℓ^+ (ℓ^-), the integrated forward/backward event ratios corresponding to the asymmetries in Fig. 1 are

$$t\bar{t}: \quad F/B = 0.95\,,$$

$$Wjjjj: \quad F/B = 1.25\,.$$

Good statistics are needed for discrimination here. The present 7 selected top quark candidate events presented by CDF [1] have ratio $F/B = 3/4$.

(b) Rank of tagged jet

The $t\bar{t} \to (b\ell\nu)(bjj)$ signal events contain two genuine b-jets; for $m_t \sim 170$ GeV these jets usually have the highest p_T and are likely to have ranks 1 and 2 (lowest rank = highest p_T). They are also by far the most likely jets to be tagged in a $t\bar{t}$ event, so the tagged jet is most likely to have rank 1 or 2 in the signal. For the $Wjjjj \to \ell\nu jjjj$ background events, however, we have argued above that by far the most likely situation is a bogus tag, with approximately equal probabilities to have rank 1, 2, 3, or 4. To illustrate these expectations, we have calculated the following probabilities $P(r)$ that the jet of rank r is tagged:

$$
\begin{array}{lllll}
t\bar{t}: & P(1) = 0.34, & P(2) = 0.33, & P(3) = 0.27, & P(4) = 0.18, \\
Wjjjj: & P(1) = 0.27, & P(2) = 0.26, & P(3) = 0.26, & P(4) = 0.24.
\end{array}
$$

These tagging probabilities have different profiles in the $t\bar{t}$ and $Wjjjj$ cases; in particular, the ratios of probabilities $P(1)/P(4)$ differ significantly:

$$
\begin{array}{ll}
t\bar{t}: & P(1)/P(4) = 1.9, \\
Wjjjj: & P(1)/P(4) = 1.1.
\end{array}
$$

The seven selected CDF top quark candidate events [1] have 3 cases with rank 1 and just one case with rank 4, giving a first estimate $P(1)/P(4) = 3$.

(c) Double-tag probability

The probabilities $P(r)$ above sum to more than 1, the excess being due to multiple tagging, mostly double-tagging. If we denote the probability that a tagged event has n tagged jets by $P(\text{n-tag})$, then $\Sigma_r P(r) = 1 + \Sigma_n (n-1) P(\text{n-tag})$. $P(\text{2-tag})$ is bigger for $t\bar{t}$ signal events (which always contain two b-jets) than for $Wjjjj$ background events (which mostly arise from bogus tags with no $b\bar{b}$ pairs present). Direct calculations with the assumed tagging efficiencies give

$$
\begin{array}{ll}
t\bar{t}: & P(\text{2-tag}) = 0.12, \\
Wjjjj: & P(\text{2-tag}) = 0.034.
\end{array}
$$

None of the 7 CDF events is double-tagged in this sense.

(d) $t\bar{t}$ invariant mass

Candidate $t\bar{t}$ events can be kinematically reconstructed, using various constraints. Equating $\not{E}_T$ to the neutrino transverse momentum, the $W \to \ell\nu$ kinematics are

determined within a two-fold ambiguity; also $W \to jj$ may be identified by the best-fitting pair of jets. The remaining two jets (one of which is tagged) are then identified as b-jets; leptonic decay information may further identify one of them as b or $\bar{b}$, but in general they are not distinguished and there are 4 possible ways of pairing them with $W \to \ell\nu$ (2 solutions) and $W \to jj$ to form the decay products of t and $\bar{t}$. We select the assignment in which the invariant masses of the two candidate top quarks agree most closely, obtaining a unique "best fit". For true $t\bar{t}$ events, this will usually give the correct assignment. For $Wjjjj$ background events, the best fit may in fact be unacceptably poor; in our analysis we have required the $W \to jj$ candidate to have $|m(jj) - M_W| < 15$ GeV and the two top quark candidate masses to agree within 50 GeV. (Note that our reconstruction method is not identical to the CDF procedure, which explicitly rescales momenta, constraining candidate W masses to equal M_W and candidate top masses to equal each other.) Background events that survive these reconstruction criteria will not necessarily agree with $t\bar{t}$ expectations in their various other kinematical distributions. Figure 2 illustrates the differences in the invariant mass $m(t\bar{t})$ distributions, which are known to be sensitive to $t\bar{t}$ production dynamics [15]; the $m(t\bar{t})$ values for the 7 reconstructed CDF events are shown by arrows. The $m(t\bar{t})$ variable is loosely related to the summed-$E_T(j)$ variable, since both reflect different aspects of the total CM energy release, but $m(t\bar{t})$ has the advantage of including the longitudinal and leptonic degrees of freedom.

(e) Lepton spectra

The lepton energy E_ℓ in the t-restframe from $t \to b\ell^+\nu$ decay is determined by the $V - A$ character of the weak couplings; in particular, it differs characteristically from the lepton spectrum in the $b' \to c\ell^-\bar{\nu}$ decay of a hypothetical fourth-generation b' quark. It is *a priori* unlikely that either of these spectra will be accurately faked by reconstructed background $Wjjjj$ events. Figure 3 illustrates the differences; the E_ℓ values for the 7 reconstructed CDF events are shown by arrows. In principle $E_\ell < \frac{1}{2}m_t$ from kinematics; the small tail in the signal at high-E_ℓ in Fig. 3 is due to momentum smearing and reconstruction errors (primarily from errors in reconstructing the longitudinal momentum of the W boson). The fourth-generation $b'\bar{b}'$ curve is shown simply for comparison; this scenario does not offer an independent explanation of the CDF events because of the low probability of tagging $b' \to cW$ decays.

We conclude that all five of the kinematical quantities (a) – (e) studied here show significant differences between the single-lepton b-tagged $t\bar{t}$ signal and the $Wjjjj$ background, and should repay further study when higher statistics are achieved.

Acknowledgments

This research was supported in part by the U.S. Department of Energy under Contract Nos. DE-AC02-76ER00881 and DE-FG03-91ER40674, by Texas National Research Laboratory Grant No. RGFY93-330, and by the University of Wisconsin Research Committee with funds granted by the Wisconsin Alumni Research Foundation.

References

[1] CDF Collaboration, F. Abe *et al.*, FERMILAB-PUB-94/097-E and Phys. Rev. Lett. **73**, 225 (1994).

[2] D0 Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **72**, 2138 (1994).

[3] For a recent discussion of the b-tagged dilepton signal, see T. Han and S. Parke, Phys. Rev. Lett. **71**, 1494 (1993).

[4] H. Baer, V. Barger, and R.J.N. Phillips, Phys. Rev. **D39**, 3310 (1989).

[5] V. Barger, J. Ohnemus, and R.J.N. Phillips, Phys. Rev. **D35**, 166 (1987); Int. J. Mod. Phys. **A4**, 617 (1989).

[6] A.D. Martin, R.G. Roberts, and W.J. Stirling, Phys. Rev. **D47**, 867 (1993).

[7] We take central values from E. Laenen, J. Smith, and W.L. van Neerven, Phys. Lett. **B321**, 254 (1994).

[8] R.K. Ellis and J.C. Sexton, Nucl Phys. **B282**, 642, (1987).

[9] V. Barger and R.J.N. Phillips, Phys. Rev. Lett. **55**, 2752 (1985).

[10] F.A. Berends, W.T. Giele, and H. Kuijf, Nucl. Phys. **B321**, 39 (1989); F.A. Berends, W.T. Giele, H. Kuijf, and B. Tausk, Nucl. Phys. **B357**, 32 (1991).

[11] V. Barger, E. Mirkes, R.J.N. Phillips, and T. Stelzer, University of Wisconsin Report No. MAD/PH/844 (1994).

[12] V. Barger, J. Ohnemus, and R.J.N. Phillips, Phys. Rev. **D48**, 3953 (1993).

[13] CDF Collaboration, F. Abe *et al.*, Phys. Rev. **D45**, 3921 (1992).

[14] F. Halzen, P. Hoyer, and C.S. Kim, Phys. Lett. **B195**, 74 (1987).

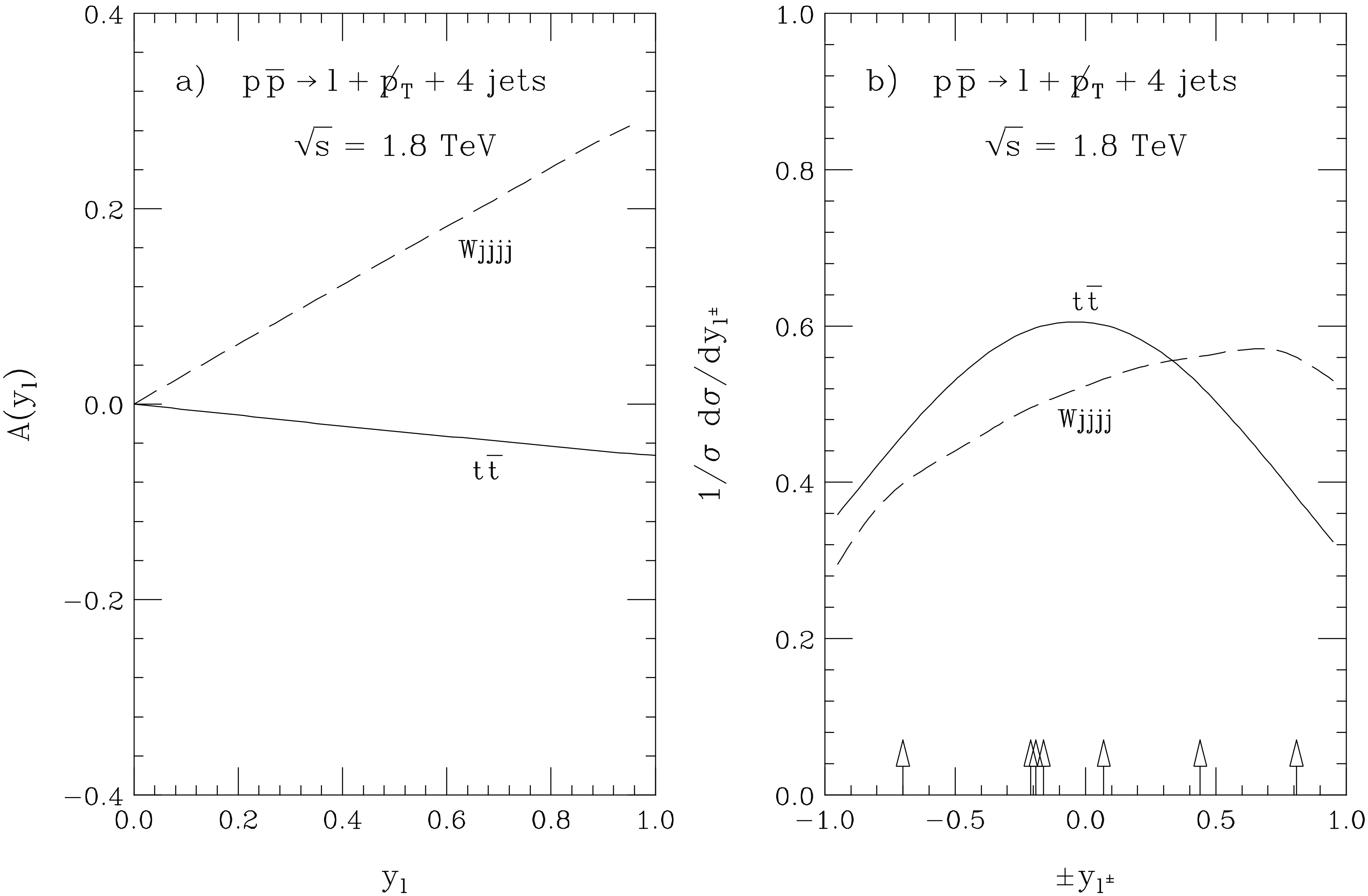
[15] C.T. Hill and S.J. Parke, Phys. Rev. **D49**, 4454 (1994); E. Eichten and K. Lane, Phys. Lett. **B327**, 129 (1994); K. Lane, Boston report BUHEP-94-25.

Figure Captions

Fig. 1. (a) Forward/backward asymmetry $A(y_\ell)$ of charged leptons versus rapidity y_ℓ for the $t\bar{t}$ signal (solid curve) and the $Wjjjj$ background (dashed curve) after all cuts. (b) The corresponding lepton rapidity distributions. The $y_{\ell^\pm}$ values for the 7 reconstructed CDF events are shown by arrows.

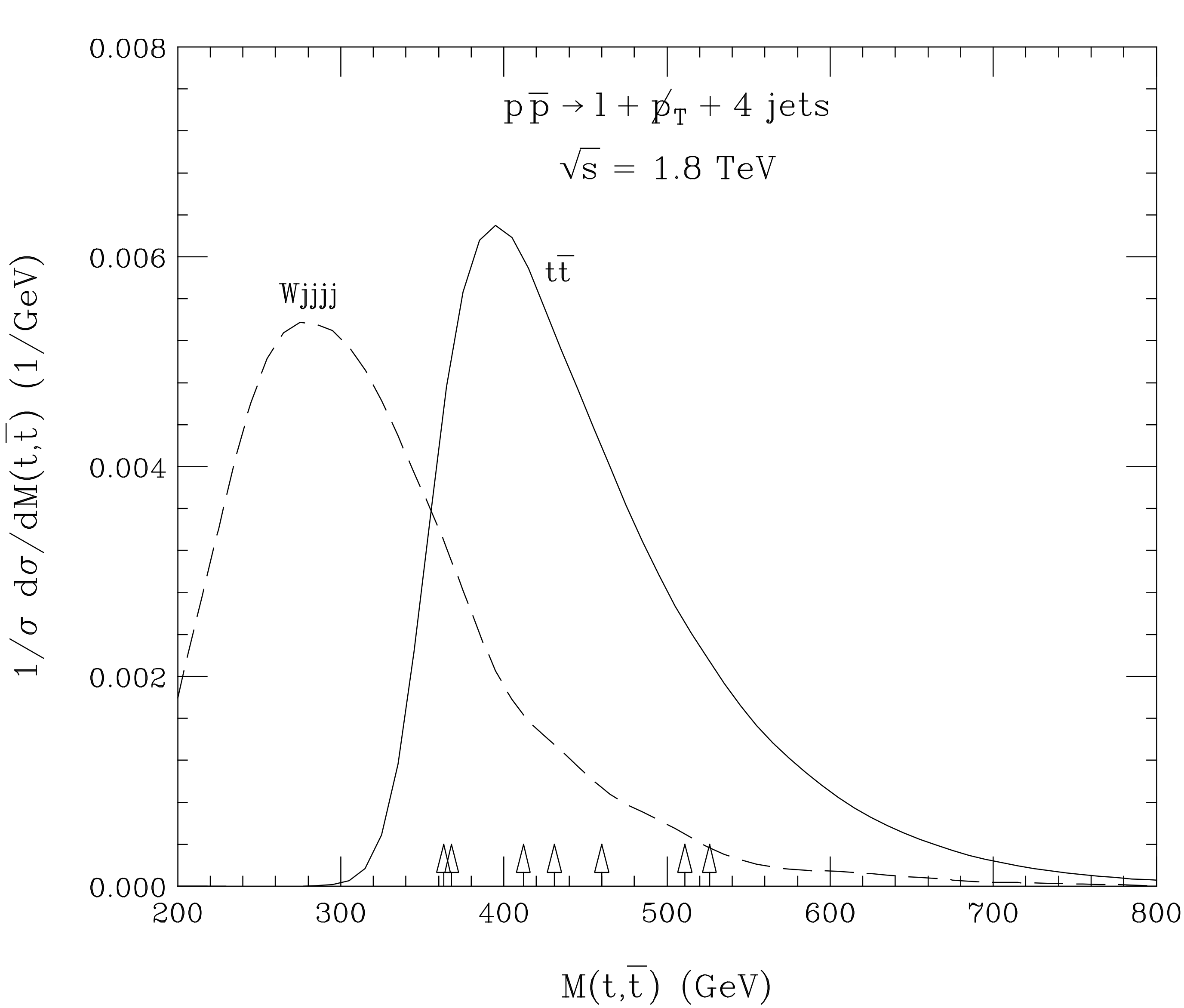
Fig. 2. Distribution of the invariant mass $m(t\bar{t}\,)$ for b-tagged reconstructed events: the solid curve denotes the $t\bar{t}$ signal and the dashed curve denotes the $Wjjjj$ background. The $m(t\bar{t}\,)$ values for the 7 reconstructed CDF events are shown by arrows.

Fig. 3. Distribution of the lepton energy E_ℓ in the parent quark restframe for b-tagged reconstructed events: the solid curve denotes the $t\bar{t}$ signal, the dashed curve denotes the $Wjjjj$ background, and the dotted curve represents contributions from hypothetical $b'\bar{b}'$ production with $m_{b'} = m_t = 174$ GeV. The E_ℓ values for the 7 reconstructed CDF events are shown by arrows. Note that $b'\bar{b}'$ events would be unlikely to be tagged.



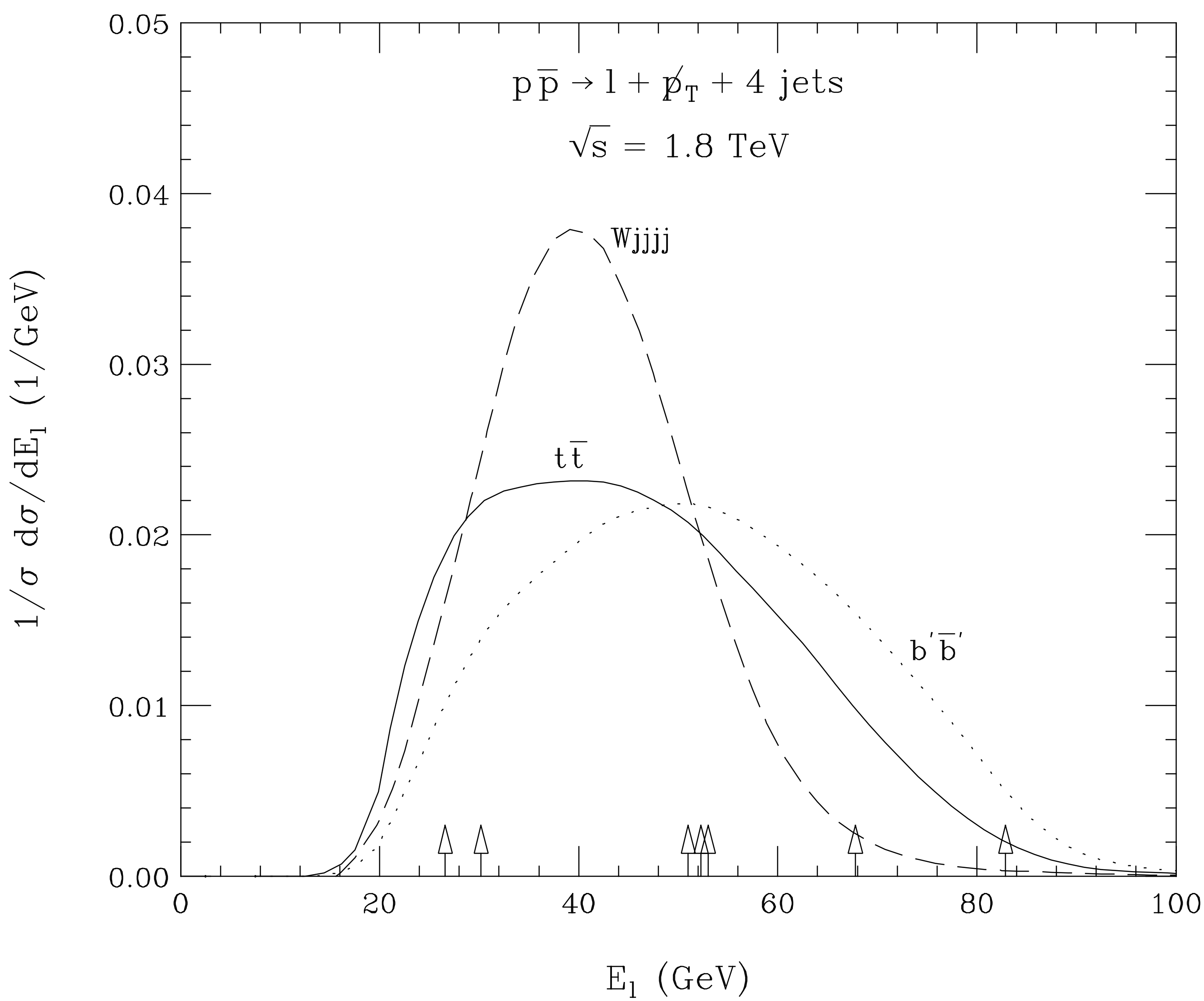
a) $p\bar{p} \to l + \not{p}_T + 4$ jets
$\sqrt{s} = 1.8$ TeV
Wjjjj
$t\bar{t}$
$A(y_l)$
y_l
0.4
0.2
0.0
−0.2
−0.4
0.0
0.2
0.4
0.6
0.8
1.0
b) $p\bar{p} \to l + \not{p}_T + 4$ jets
$\sqrt{s} = 1.8$ TeV
$t\bar{t}$
Wjjjj
$1/\sigma \; d\sigma/dy_{l^\pm}$
$\pm y_{l^\pm}$
1.0
0.8
0.6
0.4
0.2
0.0
−1.0
−0.5
0.0
0.5
1.0

Figure 1



$p\bar{p} \to l + \not{p}_T + 4$ jets
$\sqrt{s} = 1.8$ TeV
Wjjjj
$t\bar{t}$
0.008
0.006
0.004
0.002
0.000
200
300
400
500
600
700
800
$1/\sigma \; d\sigma/dM(t,\bar{t})$ (1/GeV)
$M(t,\bar{t})$ (GeV)

Figure 2



$p\bar{p} \to l + \not{p}_T + 4$ jets
$\sqrt{s} = 1.8$ TeV
Wjjjj
$t\bar{t}$
$b'\bar{b}'$
$1/\sigma\ d\sigma/dE_l$ (1/GeV)
0.00
0.01
0.02
0.03
0.04
0.05
0
20
40
60
80
100
E_l (GeV)

Figure 3